

OCENIANIE STRATEGICZNE

Daria Boratyn*
Uniwersytet Jagielloński

Streszczenie: Artykuł poświęcony jest problemowi oceniania strategicznego w systemach sędziowskich, który polega na uzyskaniu przez pewnego sędziego preferowanego przez niego wyniku wskutek podania nieuczciwej oceny lub opinii. Wychodzimy od możliwości uzupełnienia reguł ustalania werdyktu przez szkocką ławę przysięgłych w sposób uniemożliwiający ławnikom manipulowanie tym werdyktem. Badamy też odporność na manipulację werdyktem przez ławników w systemie większości sztucznej stosowanym w polskim sądownictwie. Na tej podstawie zbudowany zostaje matematyczny model systemów sędziowskich, będący uogólnieniem modelu sformułowanego przez Balinskiego i Larakiego, bliski także klasycznemu modelowi Moulina. Nowy model opiera się na założeniu, że dla dowolnych ocen lub opinii da się określić, jak bardzo są sobie bliskie. Artykuł proponuje nową definicję oceniania strategicznego, uzasadniając jednocześnie, dlaczego jest ona potrzebna. Głównym zagadnieniem pracy jest charakterystyka tych funkcji wyboru społecznego, które są w tym nowym modelu odporne na to zjawisko. Bazując na dowodzie Balinskiego i Larakiego, wykazano, że takimi funkcjami są statystyki pozycyjne, których szczególnym rodzajem jest wspomniana większość sztuczna. Artykuł zawiera również dowód twierdzenia mówiącego, że pod pewnymi dodatkowymi założeniami są to jedyne takie funkcje.

Słowa kluczowe: większość sztuczna, podejmowanie decyzji zbiorowej, ocenianie strategiczne, systemy sędziowskie, statystyki pozycyjne.

STRATEGIC GRADING

Abstract: The paper is devoted to the topic of strategic grading, which is a term that describes situations in which a judge acquires his preferred result by giving a dishonest opinion. The initial issues concern searching for the possibility of supplementing the rules of determining the verdict of the Scottish jury in a way that prevents jurors from manipulating it and strategic grading in the context of a system of artificial majority used in Polish courts of criminal justice. As it

* Daria Boratyn, Instytut Matematyki Uniwersytetu Jagiellońskiego, ul. Łojasiewicza 6, 30-348 Kraków; Centrum Badań Ilościowych nad Polityką UJ, ul. Wenecja 2, 31-117 Kraków, daria.boratyn@im.uj.edu.pl

is shown, the Polish system is immune to the strategic behaviour of the judges. Article puts forward a new model of grading systems which generalizes the model constructed by Balinski & Laraki, and is similar to the classical model of Moulin. The new model is based on the assumption that for any given grades one should be able to determine how close they are. An utterly new definition of strategic grading is given, and the article explains why it is needed. The main goal of the article is to give a characterization of grade aggregation functions which, in the new model, are immune to the strategic grading. On the basis of the proof by Balinski and Laraki, it is shown that order functions meet this criterion. Moreover, it is proven that under some additional assumptions they are the only type of such functions.

Key words: group decision making, strategic grading, grading systems, order functions.

1. PRZYPADKI SZKOCKIE

Zadaniem piętnastoosobowej szkockiej ławy przysięgłych, podobnie jak ma to miejsce w wielu innych krajach, jest orzekanie o winie lub niewinności oskarżonego. Jednakże, w przeciwieństwie do innych ław przysięgłych, szkoccy ławnicy mają do dyspozycji aż trzy możliwe werdykty: *guilty* (winny), *not guilty* (niewinny) oraz *not proven* (nie udowodniono) (Gane, 2001). W świetle prawa zarówno *not guilty*, jak i *not proven* są orzeczeniami uniewinniającymi. Jakkolwiek funkcjonalnie tożsame, werdykty te posiadają odmienne konotacje: pierwszy wyraża przekonanie ławy o niewinności oskarżonego, natomiast drugi oznacza uniewinnienie na podstawie braku wystarczających dowodów – wyrok ten obejmuje więc *szarą strefę* niepewności pomiędzy wiarą w niewinność a wymaganym dowodem winy (Bray, 2005). Stąd powszechnie uważa się, że orzeczenie *not proven* jest jakościowo gorsze dla oskarżonego od *not guilty*.

Trzeci werdykt nie jest jedyną osobliwością, z jaką musi mierzyć się kaledońska ława przysięgłych. Szkocka ustawa o postępowaniu karnym (Criminal Procedure (Scotland) Act 1995) nie precyzuje mianowicie sposobu ustalania werdyktu przez jej członków, gdy ława obraduje w pełnym piętnastoosobowym składzie. Zgodnie z wielowiekową tradycją bezwzględna większość głosów jest wystarczająca do przyjęcia dowolnego werdyktu, a konieczna dla rozstrzygnięcia skazującego (Criminal Procedure (Scotland) Act 1995; Gane, 2001). W opinii sądu apelacyjnego w sprawie *Glen v. H.M. Advocate* z 1988 roku dodatkowo podkreślono, że o ile opinia *guilty* nie uzyska poparcia co najmniej ośmiu członków ławy przysięgłych, o tyle musi ona wy-

dać werdykt uniewinniający, a także, że przekonanie jakoby każdy werdykt wymagał bezwzględnej większości głosów jest błędną interpretacją prawa (*Glen v. H.M. Advocate*, 1988). Identyczną informację zawiera również wcześniejszy precedens (*Affleck (Joseph) v. H.M. Advocate*, 1987).

Szkockie prawodawstwo nie określa jednak, jakie orzeczenie miałyby zostać wydane w przypadku, gdyby głosy przysięgłych tak się podzieliły, że żaden z werdyktów nie miałby poparcia bezwzględnej większości. Taka sytuacja miała miejsce w sprawie *Kerr (Robert) v. H.M. Advocate* z 1991 roku, kiedy głosy ławników podzieliły się w stosunku 7 do 4 do 4 oddanych odpowiednio na *guilty*, *not proven* i *not guilty*. Przysięgli nie wiedzieli, jaki werdykt mają wydać w zaistniałej sytuacji. Przewodniczący ławy przedstawił więc dokładny rozkład głosów sędziemu, który stwierdził, że w tym przypadku treścią werdyktu powinno być *not guilty* – uznał, że intencją większości ławników było uniewinnienie oskarżonego. Ostateczną wykładnię w tej sprawie zdefiniował sąd apelacyjny, który zadecydował, że w rzeczonyj sytuacji werdykt powinien brzmieć *not proven*, uzasadniając, że jest to jedyne logiczne rozwiązanie w przypadku, gdy wymagana bezwzględna większość na *guilty* nie została osiągnięta, a *not guilty* nie zdobyło względnej większości wśród głosów popierających uniewinnienie (*Kerr (Robert) v. H.M. Advocate*, 1992).

Jednakże powyższa decyzja sądu była wydana dla konkretnego rozkładu głosów i dlatego nie rozstrzyga, jak powinien wyglądać werdykt w sytuacji, gdy żadna możliwość nie ma poparcia większości bezwzględnej, ale liczby głosów za *not guilty* i *not proven* nie są równe. Niektóre jej fragmenty sugerują, że werdyktem powinno zostać to ze zdań uniewinniających, które cieszy się wśród członków ławy przysięgłych większą aprobatą. Alternatywna interpretacja mówi, że *not proven* powinno być treścią orzeczenia zawsze, gdy nie jest spełniony ani warunek większości bezwzględnej za *guilty*, ani większości zwykłej za *not guilty* (Gane, 2001). W przypadku rozkładu głosów 6 do 4 do 5 oddanych odpowiednio na *guilty*, *not proven* i *not guilty* pierwsza interpretacja prowadzi do werdyktu *not guilty*, zaś druga – do *not proven*.

W opinii wydanej w sprawie *Kerr (Robert) v. H.M. Advocate* z 1992 roku sędzia podkreślił, że w szkockiej praktyce sądowej nie leży udzielanie ławnikom dodatkowych wskazówek (w szczególności na temat tego, jak powinni zinterpretować nieoczywistą dystrybucję głosów). Zwyczajowo sprawę tę pozostawia się w gestii zdrowego rozsądku przysięgłych, co w praktyce bardzo rzadko generuje trudności z wydaniem werdyktu. Rzeczona opinia zawiera jednak informację, że jeżeli ławnicy nie byliby w stanie ustalić go między sobą i poprosili o dodatkowe instrukcje, to sędzia przewodniczący sprawie, w której ta sytuacja miałaby miejsce, powinien udzielić takiego pouczenia, które nie dawałoby im możliwości taktycznej zmiany zdania na niekorzyść oskarżonego.

Wydaje się zatem, że istnieją cztery naturalne sposoby uzupełnienia reguł ustalania werdyktu przez szkodką ławę przysięgłych o przypadek, gdy żaden z werdyktów nie ma poparcia bezwzględnej większości, a mianowicie:

- 1) orzeczenie *not guilty* bez względu na rozkład głosów pomiędzy *not guilty* i *not proven*;
- 2) orzeczenie tego z werdyktów uniewinniających, który otrzymał więcej głosów, i werdyktu *not guilty* w przypadku remisu;
- 3) orzeczenie tego z werdyktów uniewinniających, który otrzymał więcej głosów, i werdyktu *not proven* w przypadku remisu;
- 4) orzeczenie *not proven* bez względu na rozkład głosów pomiędzy *not guilty* i *not proven*.

Przeanalizujemy teraz, które z tych sposobów uniemożliwiają przysięgłym głosowanie strategiczne. Terminem tym określa się sytuacje, w których pewien członek zgromadzenia decyzyjnego (np. ławy przysięgłych) oddaje głos nieodpowiadający swoim prawdziwym preferencjom i udaje mu się w ten sposób spowodować, że kolektywna decyzja jest bliższa wynikowi, który uważa za lepszy od rezultatu, jaki miałyby miejsce, gdyby głosował zgodnie ze swoimi prawdziwymi preferencjami. Naturalne wydaje się założenie, że sędzia optujący za werdyktem *not guilty* (odpowiednio *guilty*) nie będzie przedkładał werdyktu *guilty* (odpowiednio *not guilty*) nad *not proven*.

Zacznijmy od przypadku, w którym stosujemy *regułę 1*, i rozważmy rozkład głosów członków ławy przysięgłych przedstawiony w poniższej tabeli.

	Opinia		
	<i>guilty</i>	<i>not proven</i>	<i>not guilty</i>
Liczba głosów	7	7	1

Werdykt wydany przy jej użyciu dla tego układu głosów to oczywiście *not guilty*. Wyobraźmy sobie jednak, że jeden z ławników opowiadających się za opinią *guilty* oddałby nieszczerzy głos na *not proven*. Wtedy układ głosów wyglądałby następująco:

	Zmodyfikowana opinia		
	<i>guilty</i>	<i>not proven</i>	<i>not guilty</i>
Liczba głosów	6	8	1

Werdykt w tej sytuacji brzmiałby *not proven*, gdyż opcja ta otrzymałaby bezwzględną większość głosów sędziów przysięgłych, i byłby mniej korzystny dla oskarżonego niż pierwotne *not guilty*, a więc lepszy z perspektywy ławnika przekonanego o jego winie. Oznacza to, że opisana zmiana głosu byłaby strategiczna.

Przechodząc do *reguły 2*, rozpatrzmy przedstawiony poniżej rozkład głosów członków ławy przysięgłych.

	Opinia		
	<i>guilty</i>	<i>not proven</i>	<i>not guilty</i>
Liczba głosów	5	5	5

W tym przypadku treścią werdyktu byłoby *not guilty*. Gdyby jednak którykolwiek z przysięgłych uznających oskarżonego za winnego zagłosował (nieszczercze) na *not proven*, to rozkład głosów prezentowałby się następująco:

	Zmodyfikowana opinia		
	<i>guilty</i>	<i>not proven</i>	<i>not guilty</i>
Liczba głosów	4	6	5

Werdykt uległby więc zmianie na *not proven*, co ponownie byłoby przykładem głosowania strategicznego.

Z kolei dla *reguły 3* rozważmy następujący rozkład głosów:

	Opinia		
	<i>guilty</i>	<i>not proven</i>	<i>not guilty</i>
Liczba głosów	6	4	5

W tym przypadku ponownie treścią werdyktu byłoby *not guilty*. Gdyby jednak choć jeden z ławników popierających werdykt *guilty* zagłosował (nieszczercze) na *not proven*, to rozkład głosów prezentowałby się następująco:

	Zmodyfikowana opinia		
	<i>guilty</i>	<i>not proven</i>	<i>not guilty</i>
Liczba głosów	5	5	5

Werdykt uległby więc zmianie na *not proven*, korzystniejszy dla tego ławnika. Byłby to ponowny przykład głosowania strategicznego.

Natomiast gdy stosujemy *regułę 4*, żaden z powyższych trzech przypadków zmiany głosu nie stanowi przykładu głosowania strategicznego. W istocie, sprawdziliśmy, że żaden z przedstawionych w dodatku 1 do niniejszego artykułu 28 możliwych rozkładów głosów, dla których szkockie prawodawstwo nie określa jednoznacznie wynikającego z nich werdyktu, nie daje ani jednemu z sędziów przysięgłych możliwości głosowania w sposób strategiczny.

Powyższy rezultat jest co najmniej zaskakujący, wydaje się bowiem stać w sprzeczności ze znanym **twierdzeniem Gibbarda-Satterthwaite'a** (Gibbard, 1973; Satterthwaite, 1975), mówiącym, że od problemu głosowania strategicznego nie da się uciec, gdyż każda decyzyjna funkcja wyboru społecznego określona na zbiorze co najmniej trzech mających szansę na zwycięstwo kandydatów¹ i spełniająca warunek nieograniczonej dziedziny jest albo dyktatorska, albo podatna na głosowanie strategiczne. W dalszej części artykułu zawarto wyjaśnienie tego zjawiska.

2. WPROWADZENIE DO SYSTEMÓW SĘDZIOWSKICH

Teza wspomnianego już twierdzenia Gibbarda-Satterthwaite'a dotyczy funkcji wyboru społecznego, co oznacza w szczególności, że wynik ten odnosi się do systemów wyborczych. Systemy oparte na ocenianiu, którym poświęcony jest niniejszy artykuł, mogłyby być w zasadzie traktowane jako systemy wyborcze, gdybyśmy utożsamili zbiór ocen ze zbiorem kandydatów. Jednakże w klasycznym modelu systemów wyborczych zbiór kandydatów nie jest obłożony żadnym porządkiem, natomiast każdy wyborca posiada swoją **preferencję**, czyli swój własny porządek na zbiorze kandydatów, a zbiór wszystkich możliwych preferencji jest tożsamy ze zbiorem wszystkich liniowych porządków na zbiorze kandydatów². To oznacza, że jeżeli w wyborach startuje m kandydatów, to liczba możliwych preferencji wynosi $m!$, o ile nie weźmiemy pod uwagę możliwości uznania pewnych kandydatów za równorzędnych. Dla przykładu, jeżeli w pewnych wyborach bierze udział trzech kandydatów, A , B i C , to każdy wyborca ma do wyboru jeden z następujących sześciu porządków:

$$A < B < C; A < C < B; B < A < C; B < C < A; C < A < B; C < B < A.$$

Odmierna sytuacja ma miejsce w zbiorze szkockich werdyktów, gdzie istnieje naturalny porządek:

$$not\ guilty < not\ proven < guilty.$$

Z tego powodu, tak jak wcześniej sygnalizowaliśmy, trudno wyobrazić sobie, że jakkolwiek kaledoński przysięgły mógłby preferować porządek $not\ guilty < guilty < not\ proven$ lub $not\ proven < guilty < not\ guilty$, natomiast cztery pozostałe możliwości wydają się dopuszczalne. Oznacza to, że rzeczony system nie spełnia opisanej powyżej własności niezredukowanego zbioru preferencji charakterystycznej dla systemów wyborczych. Nie jest to odosobniony przypadek. Istnieje wiele innych procedur usta-

¹ Przez mającego szansę na zwycięstwo kandydata rozumiemy takiego, dla którego istnieje co najmniej jeden układ głosów wyborców, przy którym zostaje on zwycięzcą.

² Oczywiście zdarza się, że system nie wymaga od wyborcy podania całej preferencji, a tylko jej pierwszego elementu (lub kilku pierwszych).

lania decyzji zbiorowych, które nie mogą być z tego powodu uznane za klasyczne systemy wyborcze, co w szczególności oznacza, że kwestia ich podatności lub odporności na zachowania strategiczne pozostaje otwarta. Wróćmy do tego tematu w rozdziale czwartym.

Jako pierwszy przykład rozważmy **system średniej ucinanej**. Polega on na odrzuceniu równej liczby najbardziej ekstremalnych ocen (najwyższych i najniższych), a następnie wyznaczeniu wyniku jako średniej arytmetycznej pozostałych ocen. Jest stosowany m.in. w zawodach skoków narciarskich przy ustalaniu not za styl zawodników. W konkursach sędziuje pięciu arbitrów, którzy oceniają poprawność każdego skoku w skali od 0 do 20 punktów. Dwie skrajne oceny zostają odrzucone, a notą za styl jest suma pozostałych ocen – proporcjonalna oczywiście do ich średniej arytmetycznej, zgodnie z następującym punktem III Międzynarodowego regulaminu konkursów narciarskich (*The International Ski Competition Rules Book III Ski Jumping*):

433.1 Punkty za styl

Punktacja zaczyna się od wyników pięciu arbitrów, a następnie najwyższy i najniższy z nich zostają wyeliminowane. Trzy pozostałe oceny są sumowane. Suma tych rezultatów jest ostateczną notą za styl skoku.

Nieco bardziej skomplikowanego przykładu dostarcza **system średniej skorygowanej**, który został zastosowany przez sędziów XVII Międzynarodowego Konkursu Pianistycznego im. Fryderyka Chopina w 2015 roku. Konkurs był podzielony na trzy etapy i finał. Cała procedura wyłaniania zwycięzców kolejnych etapów była dość złożona i opierała się na jednoczesnym zastosowaniu systemu punktowego i głosowania aprobującego: tak/nie. Skupimy się tutaj na procedurze punktowania zawodników. W etapach I-III każdy pianista otrzymywał od jurorów całkowitą liczbę punktów z zakresu od 1 do 25, a następnie z tych wyników obliczana była średnia arytmetyczna. Potem wszystkie oceny tego pianisty, które odbiegały od rzeczonyj średniej o więcej niż:

- 3 punkty w I etapie;
- 4 punkty w II i III etapie,

były zmieniane odpowiednio (w zależności od kierunku odchylenia od średniej) na ocenę „średnia ± 3 punkty” lub „średnia ± 2 punkty” w zależności od etapu. Przykładowo, jeżeli średnia ocen pewnego pianisty w I etapie wyniosłaby 14,35, to wszystkie jego oceny wyższe niż 17 byłyby obniżone do wartości 17,35, a wszystkie oceny niższe niż 12 byłyby podwyższone do wartości 11,35. Oceny 12, 13, 14, 15, 16 oraz 17 nie uległyby zmianie. Z tak przekształconych ocen ponownie obliczana była średnia arytmetyczna, która stanowiła wynik punktowy danego pianisty w rzeczonym etapie, stanowiący podstawę dalszego postępowania ewaluacyjnego. Powyższa procedura

została sformułowana w XIV i XV punkcie Regulaminu Jury XVII Międzynarodowego Konkursu Pianistycznego im. Fryderyka Chopina (XVII Międzynarodowy Konkurs Pianistyczny im. Fryderyka Chopina. Regulamin Jury Konkursu).

Jako ostatni przykład rozważmy **system większości sztucznej**, który jest obecnie stosowany w polskim sądownictwie. Terminem tym określa się w dogmatyce prawa taki sposób ustalania treści orzeczenia, w którym za wyrok przyjmuje się najmniej korzystną dla oskarżonego opinię spośród wystawionych przez poszczególnych orzeczników o tej własności, że ponad połowa sędziów proponuje orzeczenia nie niższe od niej. Przykładowo, jeżeli w pewnej sprawie orzeka czterech sędziów, którzy proponują skazanie oskarżonego na odpowiednio 2, 3, 4 i 5 lat pozbawienia wolności, to własność tę mają pierwsze dwie propozycje, a zatem podsądny zostanie skazany na 3 lata więzienia. Natomiast gdyby orzeczników było pięcioro, a ostatni proponowałby karę sześcioletniego pozbawienia wolności, to zasądzona zostałaby kara czteroletniego pozbawienia wolności. W Polsce rzeczony system jest stosowany w tych sprawach podlegających prawu karnemu, w których mamy do czynienia z wieloosobowymi składami orzekającymi. Ich dokładną liczebność i skład określają art. 28., art. 29. i art. 30. polskiego Kodeksu postępowania karnego (Ustawa z dnia 6 czerwca 1997 roku – Kodeks postępowania karnego). Zgodnie z wymienionymi przepisami łączna liczba sędziów i ławników orzekających w danej sprawie waha się od 1 do 5 w zależności od przedmiotu sprawy i instancji. Dokładny przebieg procedury ustalania decyzji przez wieloosobowe składy orzekające określają następujące artykuły ww. kodeksu:

Art. 109. § 1. Naradą i głosowaniem kieruje przewodniczący (...).

§ 2. Po naradzie przewodniczący zbiera głosy poczynając od najmłodszego, najpierw od ławników według ich wieku, następnie od sędziów według ich starszeństwa służbowego, a sam głosuje ostatni. Sprawozdawca, jeżeli nie jest przewodniczącym, głosuje pierwszy.

Art. 110. Narada i głosowanie nad wyrokiem odbywają się osobno co do winy i kwalifikacji prawnej czynu, co do kary, co do środków karnych, co do przepadku, co do środków kompensacyjnych oraz co do pozostałych kwestii.

Art. 111. § 1. Orzeczenia zapadają większością głosów.

§ 2. Jeżeli zdania tak się podzielą, że żadne z nich nie uzyska większości, zdanie najmniej korzystne dla oskarżonego przyłącza się do zdania najbardziej doń zbliżonego, aż do uzyskania większości.

Art. 112. Sędzia, który głosował przeciwko uznaniu oskarżonego za winnego, może wstrzymać się od głosowania nad dalszymi kwestiami; wówczas głos tego sędziego przyłącza się do zdania najprzychylniejszego dla oskarżonego.

System opisany w art. 111. to właśnie większość sztuczna. W Polsce system ten jest przenoszony bez zmian merytorycznych do kolejnych kodeksów postępowania karnego: ten zapis zawiera zarówno art. 98 ustawy z 1969 roku (Ustawa z dnia 19 kwietnia 1969 roku – Kodeks postępowania karnego), jak i art. 361 rozporządzenia z 1928 roku (Rozporządzenie Prezydenta Rzeczypospolitej z dnia 19 marca 1928 roku – Kodeks postępowania karnego). Źródłem większości sztucznej w polskim prawie jest najprawdopodobniej austriacka ustawa postępowania karnego z 1873 roku (Ustawa o postępowaniu karnym wraz z ustawą zaprowadzającą z dnia 23 maja 1873 roku), która zawiera następujące sformułowanie:

§ 20. Gdzie ustawa wyraźnie nie stanowi nic innego, wymaga się do każdej uchwały bezwzględnej większości głosów, t. j. więcej niż połowę wszystkich głosów.

Jeżeli głosy podzielone są na więcej niż dwa różne zdania tak, iż żadne z tych zdań nie ma za sobą wymaganej większości, natenczas przewodniczący będzie usiłował osiągnąć bezwzględną większość przez rozdzielenie pytań i przez powtórne zbieranie głosów.

Jeżeli i to usiłowanie pozostanie bez skutku, wówczas głosy dla obwinionego najniekorzystniejsze do głosów mniej niekorzystnych dolicza się tak długo, dopokąd się nie okaże bezwzględna większość głosów.

Wcześniejsze przypadki zastosowania większości sztucznej w prawodawstwie austriackim zawierają: § 22 ustawy dla Cesarstwa Austriackiego z 1853 roku (Ustawa o postępowaniu karnym dla Cesarstwa Austriackiego z 29 lipca 1853 roku), § 47 cesarskiego patentu z 1850 roku (Cesarski patent z dnia 17 stycznia 1850 roku – Ustawa postępowania karnego), § 425 Franciszkany (Księga ustaw na zbrodnie i ciężkie policyjne przestępstwa z 1803 roku) oraz art. 39 § 13 Terezjany (Constitutio Criminalis Theresiana z 1768 roku). Większość sztuczna była również stosowana w prawie rosyjskim, o czym świadczy art. 769 rosyjskiej ustawy postępowania karnego z 1864 roku (Ustawa postępowania karnego z dnia 20 listopada 1864 roku). Nie wygląda na to, aby jedno z tych źródeł było inspirowane drugim. Pytanie o pierwotne źródło tego przepisu pozostaje więc otwarte.

Obecnie większość sztuczna jest stosowana nie tylko w Polsce. Dotyczący jej zapis jest zawarty między innymi w: § 41 (2) austriackiego Kodeksu postępowania karnego (Strafprozeßordnung, 1975), § 216 duńskiej Ustawy o wymiarze sprawiedliwości (Bekendtgørelse af lov om retsens pleje, 2016) czy też § 36 norweskiej Ustawy o postępowaniu karnym (Lov om rettergangsmåten i straffesaker, 1981).

Rozważania na temat systemu większości sztucznej w kontekście wspomnianego wcześniej głosowania strategicznego są zawarte w kolejnym rozdziale niniejszego artykułu. W czwartym rozdziale prezentujemy natomiast ogólny matematyczny model systemów sędziowskich, czyli systemów, do których zaliczają się wymienione wyżej

przykłady. Omawiamy w nim również dotychczasowe wyniki dotyczące zachowań strategicznych w tychże systemach. Z kolei ostatni rozdział poświęcony jest przedstawieniu nowych wyników z tej dziedziny.

3. WIĘKSZOŚĆ SZTUCZNA A ZACHOWANIA STRATEGICZNE

Przyjrzyjmy się bliżej wynikom głosowań, w których decyzje ustala się za pomocą większości sztucznej. Załóżmy, że łączna liczba orzekających sędziów i ławników wynosi $n \in \mathbf{N}$. Dla ustalenia uwagi niech $0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n$ będą ich szczerymi, tj. zgodnymi z preferencjami ocenami (opiniami) uporządkowanymi w kolejności niemalejącej. Dla oskarżonego najmniej korzystna jest więc opinia α_n , a najbardziej – α_1 . Dodatkowo niech ω oznacza pierwotną treść orzeczenia, zaś $\tilde{\omega}$ – treść orzeczenia wywołaną przez zmianę oceny przez któregoś z sędziów.

Jeżeli liczba n jest nieparzysta, tzn. $n = 2k + 1$, gdzie $k \in \mathbf{N}$, to orzeczeniem zostaje środkowa opinia, czyli $\omega = \alpha_{k+1}$, ponieważ jest to najsurowsza opinia o tej własności, że wymagana większość $k + 1$ członków składu orzekającego proponuje orzeczenia nie niższe od niej. W tym przypadku wynik głosowania jest więc medianą opinii orzeczników. Jeżeli natomiast liczba członków składu orzekającego jest parzysta, czyli $n = 2k$, gdzie $k \in \mathbf{N}$, to wymagana większość również wynosi $k + 1$, a więc jako orzeczenie zostaje wybrana opinia $\omega = \alpha_k$. W tej sytuacji wynik głosowania jest więc dolną medianą opinii orzeczników.

Przeanalizujmy teraz większość sztuczną pod kątem możliwości oceniania w sposób strategiczny. Ponownie zacznijmy od sytuacji, gdy liczba orzeczników jest nieparzysta, tj. $n = 2k + 1$, gdzie $k \in \mathbf{N}$. Orzecznik, który wydał opinię równą α_{k+1} , nie mógłby odnieść żadnych korzyści ze zmiany swojego zdania, ponieważ w tym przypadku mieliśmy $\omega = \alpha_{k+1}$. Gdyby którykolwiek z sędziów i ławników wydających opinie $\alpha_i < \alpha_{k+1}$ dowolnie złagodził lub lekko zaostrzył swoje zdanie, nie miałoby to żadnego wpływu na wynik orzeczenia, tzn. $\tilde{\omega} = \omega$. Tak samo w przypadku, gdyby któryś z orzeczników wydających opinie $\alpha_i > \alpha_{k+1}$ dowolnie zaostrzył lub lekko złagodził swoją. Aby zmiana pojedynczej opinii wpłynęła na wynik głosowania, musiałaby zmieniać liczbę zdań po obu stronach α_{k+1} w porządku $\alpha_1 \leq \dots \leq \alpha_k \leq \alpha_{k+1} \leq \alpha_{k+2} \leq \dots \leq \alpha_{2k+1}$, tj. musiałaby to być zmiana z $\alpha_i < \omega$ na $\beta_i > \omega$ lub odwrotnie, z $\alpha_i > \omega$ na $\beta_i < \omega$.

W pierwszym przypadku treść orzeczenia uległaby następującej zmianie:

$$\tilde{\omega} = \begin{cases} \alpha_{k+2}, \beta_i \geq \alpha_{k+2} \\ \beta_i, \beta_i \leq \alpha_{k+2} \end{cases}.$$

Ta zmiana byłaby jednak ewidentnie niekorzystna z punktu widzenia orzecznika, który ją wywołał: jego pierwotną (szczerą) opinią było $\alpha_i < \omega$, natomiast $\tilde{\omega} \geq \omega$. W drugim przypadku natomiast nastąpiłaby poniższa zmiana treści orzeczenia:

$$\tilde{\omega} = \begin{cases} \alpha_k, \beta_i \leq \alpha_k \\ \beta_i, \beta_i \geq \alpha_k \end{cases}.$$

Ta zmiana również nie byłaby korzystna dla wywołującego ją orzecznika, gdyż jego szczerą opinią było $\alpha_i > \omega$, natomiast $\tilde{\omega} \leq \omega$.

Sytuacja parzystej liczby orzeczników, tj. $n = 2k$, gdzie $k \in \mathbb{N}$, jest całkowicie analogiczna. Ponownie zmiana zdania przez orzecznika, który wydał opinię równą α_k , nie przyniosłaby mu żadnych korzyści, gdyż w tym przypadku mieliśmy $\omega = \alpha_k$. Jedynie zmiany zdań mogące wpłynąć na zmianę treści orzeczenia to, podobnie jak w poprzednim przypadku, zmiany z $\alpha_i < \omega$ na $\beta_i > \omega$ lub odwrotnie, z $\alpha_i > \omega$ na $\beta_i < \omega$. Prowadziłyby one odpowiednio do następujących zmian treści orzeczenia:

$$\tilde{\omega} = \begin{cases} \alpha_{k+1}, \beta_i \geq \alpha_{k+1} \\ \beta_i, \beta_i \leq \alpha_{k+1} \end{cases}$$

oraz

$$\tilde{\omega} = \begin{cases} \alpha_{k-1}, \beta_i \leq \alpha_{k-1} \\ \beta_i, \beta_i \geq \alpha_{k-1} \end{cases}.$$

Analogicznie jak poprzednio, żadna z tych zmian nie byłaby więc korzystna dla wywołującego ją orzecznika.

Widzimy więc, że większość sztuczna, podobnie jak ostatnia z zaproponowanych przez nas możliwości uzupełnienia procedury ustalania werdyktu przez szcoką ławę przysięgłych, jest całkowicie odporna na strategiczne zmiany opinii przez członków składu orzekającego.

Mając to na uwadze, przyjrzyjmy się ponownie wszystkim możliwym rozkładom głosów, dla których szkockie przepisy nie dają jednoznacznego wyniku. Zakładając istnienie naturalnego porządku w zbiorze werdyktów: *not guilty* < *not proven* < *guilty*, zastosujemy regułę 4. Zamieszczona w dodatku 1 tabela zawiera informacje o tym, którą z kolei opinią w tym porządku jest opinia *not proven*, stanowiąca zgodnie z tą regułą werdykt ławy przysięgłych. Zauważmy, że w każdym z możliwych przypadków werdykt ten pokrywa się z ósmą, czyli środkową opinią, a więc jest medianą zdań wydanych przez poszczególnych ławników. Oznacza to, że odporność zarówno reguły 4, jak i większości sztucznej na zachowania strategiczne nie jest zbiegiem okoliczności, lecz wynikiem tego, że o ile składy orzekające mają nieparzystą liczbę członków, o tyle rzeczony metody ustalania werdyktów są tożsame.

Nasuwa się jednak pytanie, czy istnieją inne odporne na ocenianie strategiczne funkcje agregacji opinii. Większość sztuczna jest przykładem systemu opartego na medianie, ciekawe wydaje się więc na przykład zbadanie pod tym względem systemów opartych na średniej arytmetycznej, takich jak opisane wcześniej systemy średniej ucinanej i średniej skorygowanej. Aby formalnie rozstrzygnąć te kwestie, w kolejnym rozdziale określimy matematyczny model systemów sędziowskich.

4. MODEL SYSTEMÓW SĘDZIOWSKICH

Wymienione w niniejszym artykule systemy ustalania decyzji zbiorowych należą do klasy **systemów sędziowskich**, których matematyczny model został przedstawiony przez Michela Balinskiego i Ridę Laraki w pracy *Majority Judgment: Measuring, Ranking, and Electing* (2010; zob. też Balinski & Laraki, 2007). Poniżej opiszemy jego modyfikację.

Pierwszym niezbędnym pojęciem będzie **zbiór ocen**. Balinski i Laraki zdefiniowali go jako dowolny skończony lub będący przedziałem podzbiór $\Lambda \subset \mathbb{R}$ uporządkowany liniowo w naturalny sposób, my jednak wzbogacimy tę definicję o dodatkowy element: metrykę zgodną z tym porządkiem, mierzącą odległość ocen od siebie.

Formalnie obiekty te są zdefiniowane następująco:

Liniowym porządkiem na zbiorze Λ nazywa się dowolną relację \preceq o własnościach:

- $\alpha \preceq \alpha$ dla każdego $\alpha \in \Lambda$ (zwrotność);
- $(\alpha \preceq \beta \text{ i } \beta \preceq \gamma) \Rightarrow \alpha \preceq \gamma$ dla wszystkich $\alpha, \beta, \gamma \in \Lambda$ (przechodniość);
- $(\alpha \preceq \beta \text{ i } \beta \preceq \alpha) \Rightarrow \alpha = \beta$ dla wszystkich $\alpha, \beta \in \Lambda$ (antysymetryczność);
- $\alpha \preceq \beta$ lub $\beta \preceq \alpha$ dla wszystkich $\alpha, \beta \in \Lambda$ (spójność).

Zapis $\alpha < \beta$ oznacza, że $\alpha \preceq \beta$ i $\alpha \neq \beta$, gdzie $\alpha, \beta \in \Lambda$.

Przedziałem o początku w α i końcu w β , gdzie $\alpha, \beta \in \Lambda$ i $\alpha \preceq \beta$, nazywa się zbiór $[\alpha, \beta] = \{\gamma \in \Lambda: \alpha \preceq \gamma \preceq \beta\}$.

Metryką na zbiorze Λ jest dowolna funkcja $d: \Lambda \times \Lambda \rightarrow [0; \infty)$ spełniająca warunki:

- $d(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta$ dla wszystkich $\alpha, \beta \in \Lambda$ (oznaczoność);
- $d(\alpha, \beta) = d(\beta, \alpha)$ dla wszystkich $\alpha, \beta \in \Lambda$ (symetria);
- $d(\alpha, \gamma) \leq d(\alpha, \beta) + d(\beta, \gamma)$ dla wszystkich $\alpha, \beta, \gamma \in \Lambda$ (warunek trójkąta).

Metryka d jest **zgodna z porządkiem** \preceq , jeśli dla dowolnych $\alpha, \beta, \gamma \in \Lambda$ spełniony jest następujący warunek:

$$\alpha < \beta < \gamma \Rightarrow d(\alpha, \beta) < d(\alpha, \gamma) \text{ i } d(\beta, \gamma) < d(\alpha, \gamma),$$

tn. punkt wewnętrzny dowolnego przedziału jest oddalony od obu jego końców mniej niż te końce od siebie.

Zbiorem ocen nazywamy dowolną taką trójkę (Λ, \preceq, d) , że:

- Λ jest skończonym podzbiorem lub przedziałem w \mathbb{R} ;
- \preceq jest liniowym porządkiem na zbiorze Λ ;
- d jest metryką na Λ zgodną z porządkiem \preceq .

To właśnie arbitralny liniowy porządek i metryka, będące atrybutami zbioru ocen, stanowią kluczową różnicę między systemami sędziowskimi a klasycznymi systemami wyborczymi. Jak już wspomnieliśmy, w przypadku systemów wyborczych zbiór kandydatów nie jest wyposażony w żaden porządek ani metrykę, co sprawia, że zbiór możliwych preferencji wyborców jest równy zbiorowi wszystkich liniowych porządków na zbiorze kandydatów. Z kolei w przypadku systemów sędziowskich preferencje są indukowane przez porządek liniowy i metrykę na tym zbiorze. Konkretniej rzecz ujmując, wystarczy, że wyborca (sędzia) ustali, który kandydat (która ocena) znajduje się na pierwszym miejscu jego preferencji. Wszyscy pozostali kandydaci (oceny) zajmą w niej miejsca zgodnie z ich odległością w sensie metryki d od wybranego „faworyta”. W tej sytuacji preferencje wyborcy mają w szczególności charakter unimodalny (*single-peaked*; zob. Moulin, 1980). Oznacza to, że o ile w wyborach startuje m kandydatów, o tyle każdy wyborca ma tylko m możliwych preferencji. Przykładowo, jeżeli sędziowie mają do wyboru trzy oceny, dla ustalenia uwagi A, B i C , z zadaniem liniowym porządkiem $A < B < C$ oraz z zadaną metryką $d(A, B) = 1$, $d(A, C) = 2$, $d(B, C) = 3$, to każdy sędzia ma do wyboru jeden z trzech następujących porządków:

$$C < B < A, C < A < B, A < B < C.$$

Widzimy więc, że wzbogacenie zbioru kandydatów (ocen) o dodatkową strukturę zmniejsza liczbę możliwych preferencji wyborców (sędziów), ponieważ założony porządek nie pozwala wyborcom (sędziom) dowolnie uporządkować kandydatów (ocen). Ograniczenie dziedziny potencjalnych preferencji odróżnia przedstawiony model systemów sędziowskich od klasycznego modelu systemów wyborczych. Jest ono powodem, dla którego systemy te nie spełniają tezy (a więc i założeń) twierdzenia Gibbarda-Satterthwaite’a, albowiem gdy liczba możliwych preferencji jest zredukowana, to w szczególności istnieje potencjalnie mniej okazji dla ich strategicznych zmian. Zauważmy, że podobną cechą ma model Balinskiego i Larakiiego operujący tylko porządkiem liniowym, jednak liczba dostępnych preferencji jest w tym wypadku istotnie większa.

Warto również zauważyć, że przyjęcie bardziej ogólnego założenia o zbiorze ocen (Λ, \preceq, d) , np. że (Λ, d) jest dowolną przestrzenią metryczną spójną lub dyskretną, nie zaś podprzestrzenią \mathbb{R} , tylko pozornie prowadziłoby do szerszej definicji. Można bowiem łatwo pokazać, że struktura liniowego porządku i zgodnej z nim metryki generowałaby w tym przypadku izomorfizm między takim zbiorem ocen a odpowiadającym mu spójnym lub dyskretnym podzbiorem \mathbb{R} . Izomorfizm taki można by określić wzorem

$$\Lambda \ni \alpha \rightarrow \begin{cases} d(\alpha_0, \alpha), \alpha \geq \alpha_0 \\ -d(\alpha_0, \alpha), \alpha < \alpha_0 \end{cases}$$

gdzie α_0 jest dowolnie ustalonym elementem Λ .

Wspomnianą już kilkakrotnie funkcję przekuwającą oceny wszystkich sędziów w wynik zawodnika w konkursie sportowym, opinie członków składu orzekającego w treść orzeczenia etc., nazywa się **funkcją agregacji ocen**. Jej formalna definicja (za Balinski & Laraki, 2010) jest następująca:

Funkcją agregacji ocen jest dowolna funkcja $f: \Lambda^n \rightarrow \Lambda$ spełniająca warunki:

- $f(\alpha, \alpha, \dots, \alpha) = \alpha$ (jednomysłność);
- $f(\dots, \alpha, \dots, \beta, \dots) = f(\dots, \beta, \dots, \alpha, \dots)$ (anonimowość);
- $\alpha_j \preceq \beta_j$ dla pewnego $j = 1, \dots, n \Rightarrow$

$$f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n) \preceq f(\alpha_1, \dots, \alpha_{j-1}, \beta_j, \alpha_{j+1}, \dots, \alpha_n)$$

oraz

$$\alpha_j < \beta_j \text{ dla każdego } j = 1, \dots, n \Rightarrow$$

$$f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n) < f(\beta_1, \dots, \beta_{j-1}, \beta_j, \beta_{j+1}, \dots, \beta_n) \text{ (monotoniczność).}$$

Warunek jednomysłności dotyczy jednogłośnego podejmowania decyzji. Anonimowość gwarantuje, że zdania wszystkich sędziów mają jednakową wagę i znaczenie. Z kolei monotoniczność oznacza, że gdyby co najmniej jeden z sędziów podwyższył swoją ocenę, to wynik nie mógłby ulec obniżeniu, zaś jeśli zrobiliby tak wszyscy sędziowie, to wynik musiałby również ulec podwyższeniu. Łatwo zauważyć, że zaproponowane wcześniej *reguły 1-3* uzupełniające regułę ustalania werdyktu przez szkołą ławę przysięgłych nie są monotoniczne, a więc nie są funkcjami agregacji.

Balinski i Laraki (2010) postulowali dodatkowo ciągłość funkcji agregacji ocen. My jednak rezygnujemy z tego założenia, ponieważ wydaje się, że istnieje wiele sensownych funkcji agregacji ocen, które ciągle nie są. Dla przykładu wyobraźmy sobie konkurs akrobatyczny, w którym pewna liczba sędziów, dla ustalenia uwagi wynosząca n , ocenia występy kolejnych zawodników, przyznając im punkty z przedziału $\Lambda = [0, 50]$, przy założeniu, że sędzia ocenia występ zawodnika na 0 punktów, jeżeli

dostrzega popełnienie rażącego błędu technicznego, zaś zastosowana funkcja agregacji ocen kładzie duży nacisk na wartość techniczną występów, w związku z czym jest zdefiniowana następująco:

$$f(\alpha_1, \alpha_2, \dots, \alpha_n) = \begin{cases} \frac{\alpha_1 + \alpha_2 + \dots + \alpha_n}{n}, & \text{gdy } \alpha_1 \times \alpha_2 \times \dots \times \alpha_n \neq 0; \\ 0, & \text{gdy istnieje takie } j \in \{1, \dots, n\}, \text{ że } \alpha_j = 0. \end{cases}$$

Łatwo zauważyć, że taka funkcja, możliwa do wyobrażenia, nie byłaby ciągła.

W swoim modelu Balinski i Laraki zaproponowali następującą definicję odporności na ocenianie strategiczne:

Funkcję agregacji ocen $f: \Lambda^n \rightarrow \Lambda$ nazywamy **odporną na ocenianie strategiczne (BL)**, jeżeli dla każdego sędziego j , gdzie $j = 1, \dots, n$, i dla wszystkich $\alpha_1, \dots, \alpha_n \in \Lambda$ pełnione są warunki:

- $\alpha_j < f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n) \Rightarrow$
 $f(\alpha_1, \dots, \alpha_{j-1}, \beta_j, \alpha_{j+1}, \dots, \alpha_n) \geq f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n)$ dla każdego $\beta_j \neq \alpha_j$;
- $\alpha_j > f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n) \Rightarrow$
 $f(\alpha_1, \dots, \alpha_{j-1}, \beta_j, \alpha_{j+1}, \dots, \alpha_n) \leq f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n)$ dla każdego $\beta_j \neq \alpha_j$;

Inaczej mówiąc, funkcja agregacji ocen jest odporna na ocenianie strategiczne (BL), jeżeli sędziowie, których szczerze oceny są niższe od wartości funkcji agregacji ocen, mogą tę wartość jedynie podwyższać przez zmiany swoich ocen, a sędziowie, których oceny są wyższe od ostatecznej, mogą ten ostateczny wynik jedynie obniżyć. Ocenianiem strategicznym w sensie tej definicji jest więc każda zmiana oceny:

- niższej od wartości funkcji agregacji ocen, skutkująca spadkiem tej wartości;
- wyższej od wartości funkcji agregacji ocen, skutkująca wzrostem tej wartości.

Sformułowanie definicji oceniania strategicznego w ten właśnie sposób może wydawać się nie do końca naturalne, ponieważ uznaje ona za strategiczne również takie zmiany ocen przez sędziów, których strategiczność wydaje się być sprzeczna z intuicją. Dla przykładu wyobraźmy sobie sytuację, w której grupa sędziów ma wyrokować w sprawie młodego człowieka oskarżonego o popełnienie przestępstwa. Załóżmy, że jeden z sędziów optuje za 8-letnim wyrokiem, podczas gdy ustalona zostaje kara 7-letniego pozbawienia wolności. Wyobraźmy sobie ponadto, że wykorzystywana funkcja agregacji ocen umożliwia temu sędziemu zmianę zbiorowej decyzji na 15 lat, 20 lat lub dożywocie. W sensie przedstawionej powyżej definicji wywołanie takiej zmiany byłoby zachowaniem strategicznym, trudno jednak wyobrazić sobie, że jakikolwiek sędzia rzeczywiście uznałby taki obrót sprawy za korzystny.

Tym niemniej należy przyznać, że definicja Balinskiego i Larakiego pozwala objąć szersze sytuacje niż te, w których sędzia dąży do tego, aby ostateczna ocena składu orzekającego była najbliższa postulowanej przez niego w danej metryce. Autorzy ci zakładają bowiem, że w zbiorze ocen zadany jest tylko porządek liniowy, co dopuszcza *a priori* szerszą klasę manipulacji. Natomiast w niniejszej pracy rozważamy sytuacje, w których dysponujemy w zbiorze ocen pewną naturalną metryką.

Zanim jednak zaprezentujemy alternatywną definicję, przedstawimy twierdzenie charakteryzujące odporne na ocenianie strategiczne funkcje agregacji ocen. Balinski i Laraki udowodnili (2010; Theorem 10.1, Theorem 10.2), że jedynymi ciągłymi i odpornymi na ocenianie strategiczne funkcjami agregacji ocen w sensie definicji (BL) są **statystyki pozycyjne**, przy czym funkcję agregacji ocen $f: \Lambda^n \rightarrow \Lambda$ nazywamy **k -tą statystyką pozycyjną** (i oznaczamy dalej przez f^k) dla $k = 1, \dots, n$, jeżeli jej wartość jest równa k -tej największej ocenie, tj. $f(\gamma_1, \gamma_2, \dots, \gamma_n) = \gamma_k$, gdy $\gamma_1 \succcurlyeq \gamma_2 \succcurlyeq \dots \succcurlyeq \gamma_n$.

Większość sztuczna jest pewną szczególną statystyką pozycyjną. Mianowicie jest to $f^{\frac{n+1}{2}}$, gdy n jest nieparzyste, oraz $f^{\frac{n}{2}}$, gdy n jest parzyste. Stąd w świetle przywołanego twierdzenia odporność większości sztucznej na ocenianie strategiczne przestaje zaskakiwać. Okazuje się bowiem, że na to zjawisko odporne są w szczególności zarówno dolna mediana (czyli właśnie większość sztuczna), jak i górna mediana.

Do rozwiązania pozostaje jednak kwestia modyfikacji definicji oceniania strategicznego. Łatwo można pokazać (i zrobimy to w następnym rozdziale), że jej zawężenie nie sprawi, iż statystyki pozycyjne przestaną być odporne na to zjawisko, nietrywialna jest jednak kwestia, czy nie spowoduje ono, że do klasy opierających się próbom strategicznych zmian ocen funkcji agregacji dołączą jakieś nowe obiekty.

5. INNE SPOJRZENIE NA OCENIANIE STRATEGICZNE

Nowa definicja oceniania strategicznego, którą w tym rozdziale proponujemy, jest naturalnym przedłużeniem idei głosowania strategicznego zaproponowanej przez Gibbarda (1973) i Satterthwaite'a (1975) rozumianego jako doprowadzenie przez wyborcę do zmiany wyniku wyborów na taki, który *uważa on za lepszy*, tzn. na taki, który znajduje się wyżej na jego indywidualnej liście preferencji wyborczych. W kontekście systemów sędziowskich jako lepszą dla danego sędziego należałoby więc rozumieć każdą taką ocenę, która jest *bliżej jego szczerzej opinii*. Ocenianiem strategicznym jest więc każda taka zmiana oceny przez sędziego, która będzie skutkować zbliżeniem wartości funkcji agregacji ocen do jego szczerzej oceny. Miara tej bliskości jest natomiast wprowadzona wcześniej metryka.

Nowa definicja odporności na ocenianie strategiczne jest więc następująca:

Funkcję agregacji ocen $f: \Lambda^n \rightarrow \Lambda$, nazywamy **odporną na ocenianie strategiczne (M)**, jeżeli dla każdego sędziego j , gdzie $j = 1, \dots, n$, i dla wszystkich $\alpha_1, \dots, \alpha_n \in \Lambda$ zachodzi:

$$\min_{\beta_j \in \Lambda} \{d(f(\alpha_1, \dots, \alpha_{j-1}, \beta_j, \alpha_{j+1}, \dots, \alpha_n), \alpha_j)\} = d(f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n), \alpha_j),$$

tn. jeżeli dla każdego sędziego szczerłość jest strategią minimalizującą odległość wartości funkcji agregacji ocen od szczerzej oceny tego sędziego (choć niekoniecznie jedyną).

Wykażemy, że powyższa definicja rzeczywiście jest zawężeniem tej zaproponowanej przez Balinskiego i Larakiego (2010).

Lemat: Jeżeli funkcja agregacji ocen $f: \Lambda^n \rightarrow \Lambda$ jest odporna na ocenianie strategiczne w sensie definicji (BL), to jest odporna na ocenianie strategiczne w sensie definicji (M).

Dowód: Dla ustalenia uwagi wybierzmy dowolnego sędziego j ($j = 1, \dots, n$). Mamy do rozpatrzenia dwa nietrywialne przypadki:

$$(1) \alpha_j > f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n).$$

Wtedy, na mocy definicji (BL), dla każdej innej oceny tego sędziego $\beta_j \neq \alpha_j$ zachodzi:

$$f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) \leq f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n) < \alpha_p$$

co oznacza, że

$$d(f(\alpha_1, \dots, \beta_p, \dots, \alpha_n), \alpha_j) \geq d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j).$$

Stąd

$$\min_{\beta_j \in \Lambda} \{d(f(\alpha_1, \dots, \beta_p, \dots, \alpha_n), \alpha_j)\} = d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j).$$

$$(2) \alpha_j < f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n).$$

Wtedy, na mocy definicji (BL), dla każdej innej oceny tego sędziego $\beta_j \neq \alpha_j$ zachodzi:

$$f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) \geq f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n) > \alpha_p$$

co oznacza, że

$$d(f(\alpha_1, \dots, \beta_p, \dots, \alpha_n), \alpha_j) \geq d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j).$$

Stąd

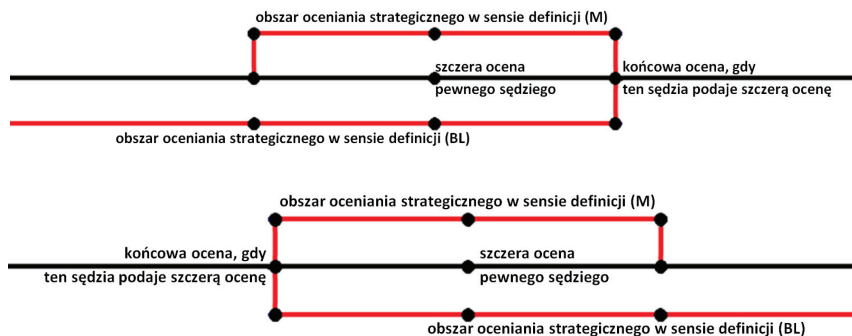
$$\min_{\beta_j \in \Lambda} \{d(f(\alpha_1, \dots, \beta_p, \dots, \alpha_n), \alpha_j)\} = d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j).$$

Zauważmy jednak, że definicje oceniania strategicznego (BL) i (M) są istotnie różne. W tym celu rozważmy następujący przykład. Załóżmy, że trzech sędziów – smakoszy – zostało zatrudnionych przez wydawcę przewodnika turystycznego do oceny nowo otwartej restauracji w pewnej miejscowości. Każdy z nich może przyznać lokalowi dowolną liczbę punktów z przedziału $[0, 100]$, zaś wykorzystywana przez wydawcę (ciągła) funkcja agregacji ocen jest następująca:

$$f(\alpha_1, \alpha_2, \alpha_3) = \begin{cases} \frac{\alpha_1 + \alpha_2 + \alpha_3}{3}, & \text{gdy } \min(\alpha_1, \alpha_2, \alpha_3) \geq \frac{2}{9}(\alpha_1 + \alpha_2 + \alpha_3), \\ \frac{3}{2} \times \min(\alpha_1, \alpha_2, \alpha_3), & \text{gdy } \min(\alpha_1, \alpha_2, \alpha_3) < \frac{2}{9}(\alpha_1 + \alpha_2 + \alpha_3). \end{cases}$$

Założmy, że sędziowie wystawiają restauracji następujące noty: 42, 60 i 78. Wtedy jej ostateczna ocena to $(42 + 60 + 78) / 3 = 60$. Gdyby jednak pierwszy sędzia zmienił swoją notę na 6, to ocena lokalu zmieniałaby się na $(3 / 2) \times 6 = 9$. Taka zmiana oceny byłaby strategiczna w sensie definicji (BL), ale już nie w sensie (M), gdybyśmy na przedziale $[0, 100]$ przyjęli metrykę euklidesową, gdyż $|60 - 42| = 18 < 33 = |42 - 9|$. Warto jednak zauważyć, że powyższa funkcja nie jest odporna na ocenianie strategiczne w sensie żadnej z definicji.

Dokładne różnice pomiędzy obiema definicjami oceniania strategicznego zobrażowano na poniższych wykresach.



Z lematu, który udowodniliśmy, wynika w szczególności, że statystyki pozycyjne rzeczywiście są odporne na ocenianie strategiczne w sensie definicji (M). Do rozstrzygnięcia pozostaje problem jedności. Udało nam się rozstrzygnąć tę kwestię przy dodatkowych założeniach spójności zbioru ocen i ciągłości funkcji agregacji ocen w formie następującego twierdzenia:

Twierdzenie: Jeżeli zbiór ocen Λ jest spójny, to statystyki pozycyjne są jedynymi ciągłymi i odpornymi na ocenianie strategiczne funkcjami agregacji ocen w sensie definicji (M).

Dowód tego twierdzenia jest treścią dodatku do niniejszego artykułu.

Zauważmy jednak, że zarówno z powyższego twierdzenia, jak i z cytowanego wcześniej rezultatu Balinskiego i Larakiiego (Balinski & Laraki, 2010; Theorem 10.1, Theorem 10.2) nie wynika odporność na ocenianie strategiczne w sensie definicji (BL) i (M) omawianych w rozdziale drugim systemów średniej ucinanej i średniej skorygowanej. Systemy te nie mieszczą się bowiem w omawianym modelu systemu sędziowskiego. W ich przypadku funkcje przypisujące ocenom sędziów ostateczne wyniki nie są funkcjami agregacji ocen w rozumieniu rozważanej przez Balinskiego i Larakiiego definicji, gdyż wyniki te mogą nie należeć do zbioru ocen, którymi dysponują sędziowie. Tym niemniej można wykazać, że systemy te nie są odporne na strategiczne zachowania sędziów.

Rozważmy najpierw system średniej ucinanej. Wyobraźmy sobie, że w trakcie konkursu skoków narciarskich pewien zawodnik miał po skoku nieudane lądowanie i otrzymał u pięciu sędziów następujące noty za styl: 13, 13, 15, 16, 17. Oceny 13 i 17 zostały odrzucone, a więc zawodnik otrzymał $13 + 15 + 16 = 44$ punkty za styl (średnia arytmetyczna tych trzech ocen wynosi wtedy 14,666...). Wyobraźmy sobie, że sędzia, który przyznał temu zawodnikowi 16 punktów, dobrze zna pozostałych arbitrów i byłby w stanie przewidzieć, że dwóch z nich oceni skok surowiej niż on (na przykład dlatego, że uznają go za podparty). Rzeczony sędzia mógłby wtedy postąpić strategicznie i zmniejszyć ich wpływ na ostateczną ocenę tego zawodnika przez przyznanie mu wyższej liczby punktów, np. 17,5. Wtedy co prawda jego nota stałaby się najwyższą i zgodnie z zasadami jako taka zostałaby odrzucona, ale ostateczna ocena skoczka i tak uległaby podwyższeniu do $13 + 15 + 17 = 45$ punktów. Średnia tych ocen wyniosłaby wtedy 15 punktów, a więc byłaby bliższa preferowanej przez sędziego oceny 16 niż wyjściowa średnia.

Z kolei w kontekście średniej skorygowanej wyobraźmy sobie, że w II etapie konkursu chopinowskiego 10 sędziów oceniło uczciwie występ pewnego pianisty, przyznając mu następujące wyniki punktowe: 12, 13, 15, 15, 16, 16, 16, 17, 18, 20. Średnia tych ocen to 15,8, a więc wyniki skorygowane to: 13,8; 13,8; 15; 15; 16; 16; 16; 17; 17,8; 17,8. Średnia skorygowana ocen tego artysty wynosiła więc 15,82. Załóżmy, że jeden z oceniających przyznałby artyście 7 punktów zamiast 12. Nowa średnia wyniosłaby wtedy 15,3, więc wyniki skorygowane byłyby następujące: 13,3; 13,3; 15; 15; 16; 16; 16; 17; 17,3; 17,3. Nowa średnia skorygowana wyniosłaby więc 15,62, czyli byłaby bliższa 12 niż wyjściowa średnia skorygowana, a zatem powyższa zmiana oceny miałaby charakter strategiczny.

Jeżeli chodzi o przypadek nieciągły, rozwiązanie zagadnienia jest problemem otwartym. Przeprowadzone symulacje sugerują jednak, że i w tym przypadku nie pojawiają się nowe funkcje agregacji ocen odporne na ocenianie strategiczne w sensie definicji (M).

6. PODSUMOWANIE

Systemy sędziowskie są szczególnym przypadkiem systemów wyborczych, w którym na zbiorze kandydatów (ocen) mamy zadany porządek. Tym terminem można więc określić szerokie spektrum sposobów ustalania zbiorowych decyzji w sprawach sądowych, zawodach sportowych i wszelakich innych konkursach, w których skład sędziowski jest wieloosobowy. Szczególnym przykładem takiego systemu jest większość sztuczna, używana w Polsce do ustalania orzeczeń w części spraw karnych. Balinski i Laraki (2010) skonstruowali model dla takich systemów, zaproponowali definicję oceniania strategicznego i udowodnili twierdzenie charakteryzujące odporne na to zjawisko funkcje agregacji ocen. W niniejszej pracy przedstawiono zmodyfikowaną wersję tego modelu, uwzględniającą istnienie naturalnej metryki w zbiorze ocen. Zaprezentowano nową, istotnie różną od wcześniejszej definicję oceniania strategicznego, nawiązującą bezpośrednio do pojęcia głosowania strategicznego w systemach wyborczych. Wykazane zostało, że statystyki pozycyjne są odporne na to zjawisko. Udowodniono również, że pod pewnymi dodatkowymi założeniami są to jedyne takie funkcje.

BIBLIOGRAFIA

- Balinski, M., Laraki, R. (2007). A theory of measuring, electing, and ranking. *Proceedings of the National Academy of Sciences of the U.S.A.*, 104, 8720–8725.
- Balinski, M., Laraki, R. (2010). *Majority Judgment. Measuring, Ranking, and Electing*. The MIT Press. Cambridge.
- Bray, S. (2005). Not Proven: Introducing a Third Verdict. *The University of Chicago Law Review*, 72, 1299–1329.
- Gane, C. (2001). The Scottish jury. *International Review of Penal Law*, 72, 259–272.
- Garde, P. (2001). The Danish jury. *International Review of Penal Law*, 72, 87–120.
- Gibbard, A. (1973). Manipulation of Voting Schemes: A General Result. *Econometrica: Journal of the Econometric Society*, 41, 587–601.
- Moulin, H. (1980). On strategy-proofness and single peakedness. *Public Choice*, 35, 437–455.
- Satterthwaite, M.A. (1975). Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. *Journal of Economic Theory*, 10, 187–217.

ŹRÓDŁA

Criminal Procedure (Scotland) Act 1995.

Affleck (Joseph) v. H.M. Advocate, 1987 S.C.C.R. 150.

Glen v. H.M. Advocate, 1988 S.L.T. 369.

Kerr (Robert) v. H.M. Advocate, 1992 S.L.T. 1031.

XVII Międzynarodowy Konkurs Pianistyczny im. Fryderyka Chopina. Regulamin Jury.

Konkursu (pol.). http://test12.nifc.pl/download/regulamin_jury_konkursu.pdf

The International Ski Competition Rules. Book III. Ski Jumping (ang.). https://assets.fis-ski.com/image/upload/v1536927329/fis-prod/assets/International_Competition_Rules_ICR_Ski_Jumping.pdf

Ustawa z dnia 6 czerwca 1997 roku – Kodeks postępowania karnego, Dz.U. 1997 nr 89 poz. 555.

Ustawa z dnia 19 kwietnia 1969 roku – Kodeks postępowania karnego, Dz.U. 1969 nr 13 poz. 96.

Rozporządzenie Prezydenta Rzeczypospolitej z dnia 19 marca 1928 roku – Kodeks postępowania karnego, Dz.U. 1928 nr 33 poz. 313.

Ustawa o postępowaniu karnym wraz z ustawą zaprowadzającą z dnia 23 maja 1873 roku, Dz.U.P. dla królestw i krajów w Radzie państwa reprezentowanych 1873 z. 42 poz. 119.

Ustawa o postępowaniu karnym dla Cesarstwa Austriackiego z 29 lipca 1853 roku, RGBl. Nr 151/1853.

Cesarski patent z dnia 17 stycznia 1850 roku. Ustawa postępowania karnego, Dz.U.P. 1850 nr 25.

Księga ustaw na zbrodnie i ciężkie policyjne przestępstwa z 1803 roku (Franciscana).

Constitutio Criminalis Theresiana z 1768 roku.

Ustawa postępowania karnego z dnia 20 listopada 1864 roku, Swod Zakonow Ros. Imp. 1912, t. 16.

Strafprozeßordnung 1975, BGBl. Nr. 631/1975.

Bekendtgørelse af lov om rettens pleje, LBK nr 1257 af 13/10/2016 (Gældende).

Lov om rettergangsmåten i straffesaker (Straffeprosessloven), 1981.

DODATEK 1

	Opinia			Werdykt
	<i>guilty</i>	<i>not proven</i>	<i>not guilty</i>	
Liczba głosów	1	7	7	2.– 8. głos
	2	6	7	3.– 8. głos
	2	7	6	3.– 9. głos
	3	5	7	4.– 8. głos
	3	6	6	4.– 9. głos
	3	7	5	4.– 10. głos
	4	4	7	5.– 8. głos
	4	5	6	5.– 9. głos
	4	6	5	5.– 10. głos
	4	7	4	5.– 11. głos
	5	3	7	6.– 8. głos
	5	4	6	6.– 9. głos

5	5	5	6. – 10. głos
5	6	4	6. – 11. głos
5	7	3	6. – 12. głos
6	2	7	7. – 8. głos
6	3	6	7. – 9. głos
6	4	5	7. – 10. głos
6	5	4	7. – 11. głos
6	6	3	7. – 12. głos
6	7	2	7. – 13. głos
7	1	7	8. głos
7	2	6	8. – 9. głos
7	3	5	8. – 10. głos
7	4	4	8. – 11. głos
7	5	3	8. – 12. głos
7	6	2	8. – 13. głos
7	7	1	8. – 14. głos

Możliwe rozkłady głosów, dla których szkodkie przepisy nie określają jednoznacznie wynikających z nich decyzji ławy przysięgłych (przy założeniu naturalnego porządku *not guilty* < *not proven* < *guilty*), wraz z werdyktami ustalonymi w oparciu o regułę 4.

DODATEK 2

Dowód: Fakt, że statystyki pozycyjne są odporne na ocenianie strategiczne w sensie definicji (M), wynika z tego, że na mocy twierdzenia sformułowanego przez Balinskiego i Larakiego (Balinski & Laraki, 2010; Theorem 10.1, Theorem 10.2) są one odporne na ocenianie strategiczne w sensie definicji (BL), a więc, korzystając z lematu zamieszczonego w rozdziale piątym, także w sensie definicji (M).

Pozostaje udowodnić jedyność. W tym celu wystarczy wykazać, że każda ciągła funkcja agregacji ocen odporna na ocenianie strategiczne w sensie definicji (M) jest też odporna na ocenianie strategiczne w sensie definicji (BL), gdyż wówczas będziemy mogli skorzystać z przywołanego wyżej twierdzenia Balinskiego i Larakiego. Dowód przeprowadzimy metodą nie wprost.

Niech $f: \Lambda^n \rightarrow \Lambda$ będzie odporną na ocenianie strategiczne w sensie definicji (M) funkcją agregacji ocen, dla której istnieją: sędzia $j \in \{1, \dots, n\}$, układ ocen $(\alpha_1, \dots, \alpha_n)$ i różna od α_j ocena β_j , takie że

$$\beta_j < \alpha_j < f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n) \text{ i } f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) < f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n)$$

lub

$$\beta_j > \alpha_j > f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n) \text{ i } f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) > f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n).$$

Dla ustalenia uwagi załóżmy, że spełniony jest pierwszy z powyższych warunków.

Pokażemy, że $f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) < \alpha_j$.

Z $\alpha_j < f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n)$ wynika, że $d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j) > 0$. Następnie z odpor-
ności funkcji f na ocenianie strategiczne w sensie definicji (M) otrzymujemy

$$d(f(\alpha_1, \dots, \beta_p, \dots, \alpha_n), \alpha_j) \geq d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j) > 0,$$

czyli $f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) \neq \alpha_j$. Załóżmy więc nie wprost, że $f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) > \alpha_j$. Na
mocy monotoniczności funkcji f spełniona jest więc nierówność

$$\alpha_j < f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) < f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n),$$

co ze względu na zgodność metryki z porządkiem oznacza, że

$$d(f(\alpha_1, \dots, \beta_p, \dots, \alpha_n), \alpha_j) < d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j),$$

i otrzymujemy sprzeczność.

Wówczas

$$f(\alpha_1, \dots, \beta_p, \dots, \alpha_n) < \alpha_j < f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n).$$

Z faktu, że Λ jest przedziałem, wynika, że zbiór

$$\{\alpha_1\} \times \dots \times \{\alpha_{j-1}\} \times [\beta_p, \alpha_j] \times \{\alpha_{j+1}\} \times \dots \times \{\alpha_n\} \subset \Lambda^n$$

jest spójny jako iloczyn kartezjański skończonej liczby podzbiorów spójnych Λ . Stąd
wynika, że zbiór

$$f(\alpha_1, \dots, \alpha_{j-1}, [\beta_p, \alpha_j], \alpha_{j+1}, \dots, \alpha_n)$$

jest spójny w Λ , gdyż obraz zbioru spójnego przez funkcję ciągłą zawsze jest spójny.
Jednocześnie

$$\alpha_j \notin f(\alpha_1, \dots, \alpha_{j-1}, [\beta_p, \alpha_j], \alpha_{j+1}, \dots, \alpha_n)$$

ponieważ gdyby istniało takie $\gamma_j \in \Lambda$, że $f(\alpha_1, \dots, \gamma_p, \dots, \alpha_n) = \alpha_j$ to na mocy mono-
toniczności funkcji f spełnione byłyby następujące nierówności:

$$\beta_j < \gamma_j < \alpha_j$$

oraz

$$d(f(\alpha_1, \dots, \gamma_p, \dots, \alpha_n), \alpha_j) = 0 < d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j),$$

a to jest niemożliwe, ponieważ korzystając z odporności funkcji f na ocenianie strategiczne w sensie definicji (M), mamy:

$$d(f(\alpha_1, \dots, \alpha_p, \dots, \alpha_n), \alpha_j) = \min_{\delta_j \in \Lambda} \{d(f(\alpha_1, \dots, \delta_p, \dots, \alpha_n), \alpha_j)\}.$$

Z drugiej strony, jak pokazaliśmy wyżej,

$$\alpha_j \in [f(\alpha_1, \dots, \alpha_{j-1}, \beta_j, \alpha_{j+1}, \dots, \alpha_n), f(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \dots, \alpha_n)].$$

Wynika z tego, że $f(\alpha_1, \dots, \alpha_{j-1}, [\beta_j, \alpha_j], \alpha_{j+1}, \dots, \alpha_n)$ nie jest spójnym podzbiorem Λ , co oznacza sprzeczność.